CELL NUCLEUS SEGMENTATION OF SKIN TUMOR USING IMAGE PROCESSING

Toshiyuki Tanaka[1], Tomoo Joke[1], Teruaki Oka[2]

[1]Faculty of Science and Technology, Keio University, Yokohama, Japan
[2]Division of Pathology and Clinical Laboratory, Kanto Central Hospital, Tokyo, Japan

Abstract-**Automation and quantification of diagnosis of tumor cell images have been studied for these three decades in the field of medical imaging technology. Many techniques of image processing were proposed to solve problems such as nucleus segmentation and classification. But these studies have mainly focused on epithelial tumors. Nonepithelial skin tumors such as dermatofibroma (DF) and dermatofibrosarcoma protuberans (DFSP) have not been enough studied. DF is benign tumorous disease and DFSP is mid-grade malignant tumor. Recently, it is necessary that criterion of classification between DF and DFSP is quantitatively specified. In this paper a system for segmenting cell nuclei of DF and DFSP is proposed. Nuclei regions are objectively segmented and surrounded using edges of strength by the system. Segmentation of arbitrary shaped nuclear regions and weakly stained nuclear region is made. A dynamic thresholding method with combining Laplacian histogram with Ohtsu's method is used for segmentation. Segmentation test was done using real tissue cell images of DF and DFSP to evaluate validity of this system. Shape characteristics such as grade of similarity to circle were also computed from the segmented regions to assure that some differences between DF and DFSP is expressed in its distribution.**

Keywords – **Segmentation, skin tumor, cell nucleus, dynamic threshold, Laplacian histogram**

I. INTRODUCTION

Automation and quantification of diagnosis of tumor cell image have been studied for these three decades in the field of medical imaging technology. Most of studies are development of diagnosis support system that solved problem of pathologist shortage. One of the studies is Pap semear screening systems for cervical cancer that have already released[1][2]. Many techniques of image processing were proposed to handle with problems such as nucleus segmentation and classification in development of these systems[3][4]. But these studies mainly focus on epithelial tumors, and there are few studies that focus on nonepithelial tumors. Dermatofibroma (DF) and deamatofibrosarcoma protuberance (DFSP) are nonepithelial skin tumors that need be distinguished. DF is a benign tumorous disease and DFSP is mid-grade malignant tumor. Recently, it was reported that probability of spectrum of DF was different form that of DFSP and criterion of classifying these two tumors is continually discussed[5]. An important criterion of tumor classification is morphological characteristics of tissue cell image. Although distribution of cell nucleus area and shape characteristics such as grade of circle are mainly used in the previous reports, there are some problems in terms of the amount of work and objectivity if nucleus segmentation is done manually or semiautomatically. There are a lot of ambiguous nuclei areas in tissue cell images of DF and DFSP. Therefore the problem of objectivity is not ignorable. It is difficult to segment nuclei regions accurately, because there
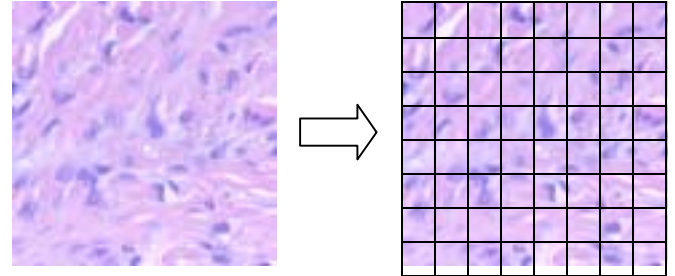

Fig.1 Division of original sarcoma image into 64 small regions that have the same areas.

exists a lot of ambiguous nuclei as already stated. In this paper an automatic system to segment cell nuclei of DF and DFSP is proposed. Nuclei regions are objectively segmented and surrounded using strength of edges by the system. Segmentation of arbitrary shaped nuclear regions and weakly stained nuclear region is made. A dynamic thresholding method[6][7] with combining Laplacian histogram method[8] with Otsu's thresholding method[9] is used for segmentation.

II. METHODOLOGY

Here we explain contour extraction method of cell nucleus of skin tumor. This method bases on Laplacian histogram and ohtsu's thresholding method. We used digital images for this study, which were taken into computer by microscope camera with magnification of 40 and filmscanner. Those images have size of 800× 500 pixels. Tumors in the image are dyed using Hematoxylin-Eosin (HE) staining method. Two images with 400× 400 pixels are taken from the original image and they are used as experimental samples. 20 images of dermatofibroma and 20 images of dermatofibrosarcoma are used in this study.

Step 1: First, we transform the original mage with RGB bases into the image with YIQ bases, using the following equation.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.522 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (1)$$

In the above equation R, G, B means values of RGB bases in each pixel of image, respectively. Component Y is usually used for binarization in general image processing. But it was shown from our research that component R in RGB bases was more adequate for binarization of tumor images, since nuclei are dyed to violet by HE dyeing method. Component R is used for binarization in this study.

Step 2: Single threshold is usually used for binarization of image. Since intensity is uneven in some parts of tumor image, single threshold reaches in inadequate result for the tumor image. Binarization is done under such a condition by

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
15 Oct 2001	N/A	-

Title and Subtitle	Contract Number
Cell Nucleus Segmentation of Skin Tumor Using Image Processing	Grant Number
	Program Element Number

Author(s)	Project Number
	Task Number
	Work Unit Number

Performing Organization Name(s) and Address(es)	Performing Organization Report Number
Faculty of Science and Technlogy Keio University Yokohama, Japan	

Sponsoring/Monitoring Agency Name(s) and Address(es)	Sponsor/Monitor's Acronym(s)
US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.

Abstract

Subject Terms

Report Classification	Classification of this page
unclassified	unclassified

Classification of Abstract	Limitation of Abstract
unclassified	UU

Number of Pages
4

making threshold surface. We divide the image into 64 small regions with 50×50 pixels as shown in Fig.1 for making the threshold surface.

Step 3: Absolute values of Laplacian are generally large near the edge of objects in gray scale image. Gray scale image means the image of R component in this study. Laplacian histogram is computed for each divided region of component R. Let $f(x, y)$ be digital image of tumor. Ordered pair (x, y) means coordinates of reference pixel. Laplacian is computed using the following equation.

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} = f_{xx}(x,y) + f_{yy}(x,y)$$

where

$$
\begin{aligned}
f_{xx}(i,j) &= \{f(i+1,j) - f(i,j)\} - \{f(i,j) - f(i-1,j)\} \\
&= f(i+1,j) - 2 \cdot f(i,j) + f(i-1,j) \\
f_{yy}(i,j) &= \{f(i,j+1) - f(i,j)\} - \{f(i,j) - f(i,j-1)\} \\
&= f(i,j+1) - 2 \cdot f(i,j) + f(i,j-1)
\end{aligned}
\tag{1}
$$

After absolute value of Laplacian is computed for each pixel of image data, 10% of all the pixels are selected in numerical order from the maximum value.

Step 4: Next, threshold of binarization for each divided regions are obtained using Ohtsu's method. When maximum gray scale level is L, intensity range is $S = \{1, 2, \cdots, L\}$, pixel number with level i is n_i and number of all the pixels is $N = n_1 + n_2 + \cdots + n_L$, intensity histogram is computed by the following equation.

$$p_i = \frac{n_i}{N} \quad \left(i \in S, p_i \geq 0, \sum_{i=1}^{L} p_i = 1 \right)$$

When k is threshold in this region, we separate the whole region of gray scale into two classes, one is $S_1 = \{1, 2, \cdots, k\}$ and another is $S_2 = \{k+1, \cdots, L\}$. Occurrence probability ω_1, ω_2 of each classis is computed by the following equation.

$$\omega_1 = \sum_{i=1}^{k} p_i, \ \omega_2 = \sum_{i=k+1}^{L} p_i = 1 - \omega_1 \tag{2}$$

Mean intensity μ_1, μ_2 of each class is obtained by the next equation.

$$\mu_1 = \sum_{i=1}^{k} i p_i / \omega_1, \ \mu_2 = \sum_{i=k+1}^{L} i p_i / \omega_2 \tag{3}$$

Mean intensity μ_T of the whole class is computed by the next equation.

$$\mu_T = \sum_{i=1}^{L} i p_i \tag{4}$$

Valiance σ_1, σ_2 of each class is computed by the following equation.

$$\sigma_1^2 = \sum_{i=1}^{k} (i - \mu_1)^2 p_i / \omega_1, \ \sigma_2^2 = \sum_{i=k+1}^{L} (i - \mu_2)^2 p_i / \omega_2 \tag{5}$$

Criterion η is used for obtaining the optimal threshold k.

$$\eta = \frac{\sigma_B}{\sigma_T} \tag{6}$$

where

$$\sigma_T^2 = \sum_{i=1}^{k} (i - \mu_T)^2 p_i, \ \sigma_B^2 = \omega_1 \omega_2 (\mu_1 - \mu_2)^2 \tag{7}$$

σ_T in the above equation means valiance of the whole class and σ_B means valiance between each class.

Step 5: Here threshold surface is made as shown in Fig.2 using threshold of each small region obtained in step 4. Coordinate in parameter plane (s, t) is assumed in the following equation.

$$x = x(s,t), \ y = y(s,t), \ z = z(s,t) \tag{8}$$

These coordinates are also expressed by B-spline function $B_{m,n}(x)$ as below.

$$x(s,t) = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \alpha_{ij} B_{i,K}(s) B_{j,L}(t) \tag{9}$$

$$y(s,t) = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \beta_{ij} B_{i,K}(s) B_{j,L}(t) \tag{10}$$

$$z(s,t) = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \gamma_{ij} B_{i,K}(s) B_{j,L}(t) \tag{11}$$

Threshold surface is obtained by making a set of equations from thresholds of each divided regions and by solving $\alpha_{ij}, \beta_{ij}, \gamma_{ij}$ in the above equation.

Step 6: Binarization is done to each pixel by using threshold surface obtained in step 5.
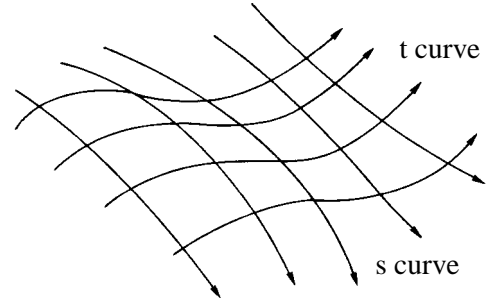

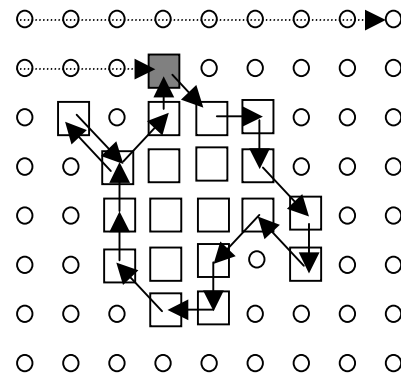Fig.2 Threshold plane obtained by B-spline function
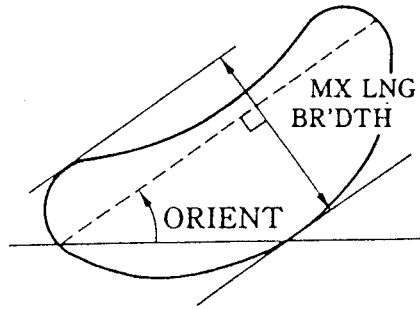

Fig.3 Tracking the contour points

Fig.4 Grade of needle

Step 7: Three processes are done as postprocessing. First processing is to eliminate the nuclei regions on the edge of image. Second processing is to eliminate intranuclear space, which is background region within the nuclei region. Final processing is to expand and constrict the obtained nuclei regions for smoothing the edges of nuclei.

Step 8: Many shapes of nuclei are obtained in an image after binarization. Since we must obtain feature of each nucleus, labeling is done for each nucleus region. Different numbers are labeled nuclei in image in labeling process. After labeling process each shape of nuclei can be processed in order.

Step 9: Contour tracking is made as shown in Fig.3. Starting point of tracking is scanned from the left upper of digital image to the right lower. After starting point is found, contour points are found out by searching clockwise on 8 neighborhoods. Tracking the contour will be finished when noticed point returns to starting point.

Step 10: Area, perimeter, grade of circle and grade of needle are selected as features of nuclei. Grade of circle (C) is obtained by the next equation, when L is perimeter and A is area of nucleus.

$$C = 4\pi A / L^2 \tag{12}$$

Grade of needle (SF)[10] is computed by the following equation.

$$SF = \frac{MX\ LNG}{BR'\ DTH} \tag{13}$$

MX LNG and BR'DTH in (13) are shown in Fig.4.
We can know of polymorphism and atypism of nulei from those features of contours.

III. RESULTS

20 dermatofibroma and 20 dermatofibrosarcoma with magnification of 40 were used for this research. An example of contour extraction for dermatofibroma is shown in Fig.5 and that of dermatofibrosarcoma is shown in Fig.6. Histograms of areas of nuclei were obtained from results of contour extraction as shown in Fig.7 and Fig.8. Fig.7 shows histogram of dermatofibroma and Fig.8 shows that of dermatofibrosarcoma. Frequency distributions have similar shapes each other, but cumulative frequencies are completely different. This result is the reason that average number of nuclei per one image for dermatofibrosarcoma is much more than that for dermatofibroma. Histograms of grade of circle and grade of needle are shown in Fig.9 and Fig.10. These figures show histograms of dermatofibroma. Also in the case
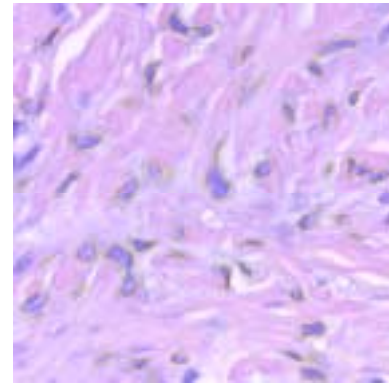

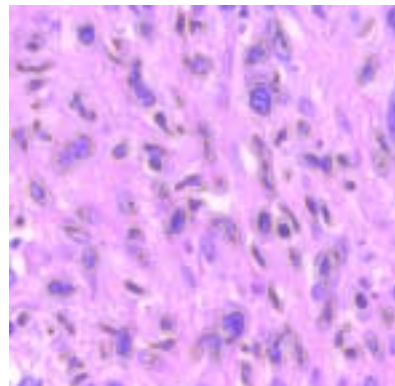Fig.5 Contour extraction of dermatofibroma

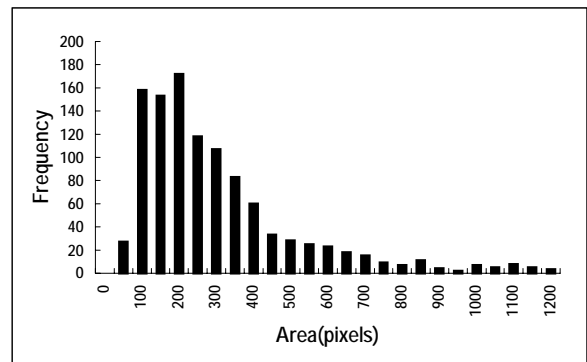Fig.6 Contour extraction of dermatofibrosarcoma

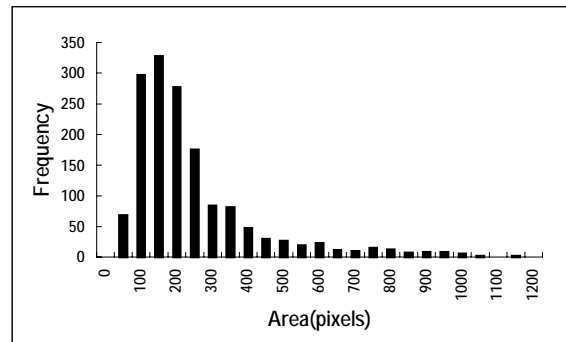Fig.7 Histogram of area of nuclei (dermatofibroma)

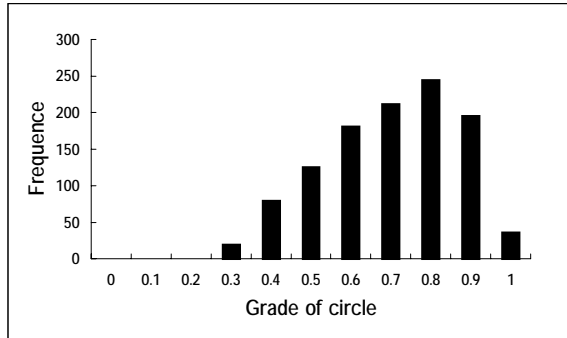Fig.8 Histogram of area of nuclei (dermatofibrosarcoma)

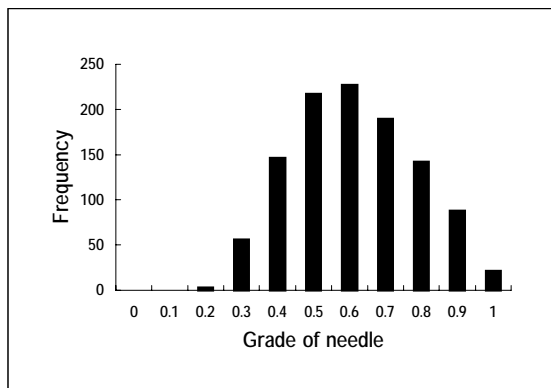Fig.9 Histogram of grade of circle (dermatofibroma)

Fig.10 Histogram of grade of needle (dermatofibroma)

of grade of circle and grade of needle, frequency distributions have similar shapes and cumulative frequencies are completely different. Although we have not computed the average and variance and other statistics yet because of shortage of image data, it seems that difference between dermatofibroma and dermatofibrosarcoma appears in the obtained results.

IV. DISCUSSION

Contour extraction by the proposed method is good enough to obtain shapes of nuclei of tumors. When nuclei have indistinct edge, the nuclei are ignored from contour extraction. Even if nuclei have indistinct edge, when there is obvious region within nuclei, the edge of obvious region is obtained. The matter that is dyed to violet in nucleus by HE dyeing method is called chromatin. The proposed method is objective from the viewpoint of distribution of chromatin. Nuclei with less chromatin have weak and indistinct edge.

When more than two nuclei superimpose each other, since the proposed method regard the connected region as one nucleus, one large contour is obtained.

Area, grade of circle and grade of needle are used as index of nuclei shape. Although the significance between malignant tumor and benign tumor appeared in obtained results, it is not enough for doctors to diagnose dermatofibroma and dermatofibrosarcoma with this system. We must expand this method to that with high accuracy from diagnostic viewpoints.

V. CONCLUSION

This paper proposed system with quantitative criterion for classification between DF and DFSP. First, after discussion about previous diagnosis of uterine cell, we made a system that consists of segmentation of divided image, extraction of nuclei region by combining Lapracian histogram with Ohtsu's method. Next, experiment of extraction was done by the system when cytohistological image of DF and DFSP were used as reference samples. It was confirmed that the proposed system could extract regions of arbitrary shape with obvious edge. The whole shapes were extracted for nuclei with distinct edge and only part of chromatin concentrated regions were extracted for nuclei with blurred edge. Shape characteristics such as grade of similarity to circle were also computed from the segmented regions to confirm some differences between DF and DFSP. Segmentation test was done using real tissue cell images of DF and DFSP. Distribution of shapes characteristics such as area and grade of similarity to circle for DF and DFSP showed validity of quantitative criteria of the system from the results of segmentation test. The proposed method is adequate for medical study and diagnosis support of pathologists.

REFERENCES

[1] Pascal Bamford,Brian Lovell, "Unsupervised cell nucleus segmentation with active contours," Signal Processing, Vol.71, pp203-213, 1998

[2] Laurence O'Gorman, Arthur C.Sanderson,Kendall PrestonJr," Image segmentation and nucleus classification for automated tissue section analysis," Proceedings of the IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition, pp.89-94, 1983

[3] W.E.Blanz, E.R.R.Reinhardt"Image Segmentation by Pixel Classification," Pattern Recognition, Vol.13, No.4, pp.293-298, 1981

[4] L.Vincent, P.Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.13, pp.583-598,1991

[5] Horenstein MG, Prieto VG, Nuckols JD," Indeterminate fibrohistiocytic lesions of the skin: is there a spectrum between dermatofibroma and dermatofibrosarcoma protuberans?," American journal of surgical pathology, Vol.24, pp996-1003, 2000

[6] C.K.Chow, T.Kaneko,"Automatic boundary detection of the left ventricle from cineangiograms," Comput. Biomed.Res.5,pp388-410,1972

[7] Yasuo Nakagawa, Azriel Rosenfeld," Some Experiments on variable thresholding," Pattern Recognition, Vol.11, pp.191-204, 1979

[8] J.S.Weszka,R.N.Nagel,A Rosenfeld, " A Threshold Selection Technique," IEEE Trans. Computer, pp.1322-1326, 1974

[9] N.Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. System, Man, and Cybernetics, Vol.SMC-9, No.1, pp.62-66, 1979

[10] E.Bengtsson," The measuring of cell features," Analytical and quantitative cytology and histology, Vol.9, No.3, pp212-217, 1987